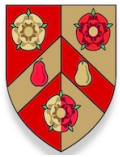
Momentum, mean reversion and meeting intensity in cryptocurrency market

Zi Yin

Wolfson College



University of Oxford

Word count: 12467

Submitted in partial fulfilment of the requirements of the degree of

Master of Philosophy in Economics

Trinity 2020

Acknowledgements

Firstly, I would like to thank my supervisor, Dr Matthias Qian, for his guidance during my study at the University of Oxford. My sincere thanks also go to Dr Xiyu Jiao, Dr Jan-Peter Calliess, and Dr Yang Wang for their encouragement, insightful comments, and enlightenment at first glance of this research. Secondly, I wish to thank my mother, Zhengzheng Yin, grandmother, Enzhu Wang, aunt, Jianjian Yin, and brother, Jiayan Wang, for supporting me in my academic endeavours. Thirdly, I would like to express my deep gratitude to my friends in Oxford: Dr Siu Yiu, Xuan Li, Yifan Zhang, Dr Kai Zhang, Dr Thomas Tai-Dung Chen and Anger Hang, for taking part in the useful decision and accompanying me through the hard times; Edward Moyse and Louise Croft for providing financial supports and giving precious guidance to my future career; Cynthia Chen, Jenny Liu, Qiqi Dong, Xiaoting Ma, Teng Zhou, Juanyi Li, Jing Wang and Yifang Cheng for calling me internationally and supporting me both on and off the water. Finally, I would like to thank Xinxin Zhang and Lidingrong Huang who volunteered their time to proofread this thesis.

Declaration

I, Zi Yin, declare that the work presented in this MPhil thesis is, to the best of my knowledge and belief, original and my own work, except as acknowledged in the text, and that material has not been submitted, either in whole or in part, for a degree at this or any other university.

Zi Yin	*10 June 2020*
Signature	*Date*

Abstract

Andrei and Cujean (2017) have proposed a joint theory of the return momentum and reversal based on a rational-expectation model under the information percolation framework. A legitimate question is what empirical exercise could validate their model. This thesis tests their model with word-of-mouth communication by using high-frequency data in cryptocurrency markets. I show evidence that time-series momentum reflects the speed of asset-specific information percolation. I establish two key results in the cryptocurrency market. Firstly, cryptocurrency returns exhibit short-term momentum and long-term reversal. Secondly and more importantly, both of the magnitude of the short-term momentum effect and long-term reversal effect increase with the meeting intensity measured by the Google trend index. These findings are consistent with the theoretical predictions derived by Andrei and Cujean.

Table of Contents

List of Figures

List of Tables

1 | Introduction

1.1 Background

Price momentum effect has been widely documented across different asset classes and markets. The literature distinguishes between two categories of momentum effects: cross-sectional momentum and time-series momentum. The cross-sectional momentum focuses on the relative performance of securities in different sections. A financial asset exhibits cross-sectional momentum if its return keeps outperforming/underperforming its peers' returns consistently (Jegadeesh and Titman, 2001; Hong et al., 2000). The time-series momentum, also the centre of this thesis, focuses purely on a security's own past performance. An asset exhibits time series momentum if its return shows strong positive predictability from its own past returns - past winners continue to perform well, and past losers continue to perform badly. Various theories, mostly behavioral, have been proposed in an effort to explain the source of both types of momentum (Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999). Apart from the standalone momentum phenomenon, researchers also find that a phase of return reversal over a longer horizon often occurs in most of the financial markets. For instance, Moskowitz et al. (2012), using the data of 58 liquid instruments across all kinds of financial markets, have documented that the magnitude of price momentum is large for short look-back periods (one to six months) and decays as the look-back period increases, with additional evidence of return reversal for time horizons longer than a year.

Although the co-existence of the short-term momentum and long-term reversal effects in financial market does not sound controversial, it is much more difficult

to derive a coherent theory to explain both of them. It is challenging to explain the persistence of these two effects because if the professional investors know the return patterns, they will take advantage to trade on them, resulting in the elimination of the patterns. Conrad and Kaul (1998) propose a risk-based interpretation of momentum but there is little empirical evidence for their risk theory. Apart from the lack of empirical support, the momentum effect is also not subsumed in the three-factor risk model by Fama and French (1996). As for the non-risk based explanations, mostly behavioral theories, are able to give rise to positive short-term return autocorrelations and long-term reversal by assuming traders possessing different behavioral bias (Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999). However, the empirical evidences that show the link between various measures of investor sentiment and momentum measures are not very strong. For example, Moskowitz et al. (2012) indicate that behavioral models have yet to identify the main source driving momentum and reversal. Fortunately, Andrei and Cujean (2017) solve this problem by proposing a theory based on a rational expectation model (Grossman and Stiglitz, 1980) under the information percolation framework (Duffie and Manso, 2007). They build a model of centralized trading based on a noisy rational-expectations equilibrium with decentralized information gathering achieved by applying the information percolation process. In their model investors can randomly meet with each other and share information about the asset they trade in bilateral private meetings. As agents accumulate information through random meetings, the average precision of information in the economy increases at an accelerated rate. The increase in market precision, after a certain condition is satisfied[1], could generate short-term momentum and long-term reversal in returns despite the fact that investors could trade on them. This is because through the meeting process, agents acquire heterogeneous amounts of

[1]They prove that beyond a certain threshold of the frequency at which agents meet each other, their theory can generate both short-term momentum and long-term reversal in returns.

information, leading them to implement different trading strategies. More specifically, they show that better-informed investors systematically trade against the momentum, on the contrary, less-informed investors ride the momentum. In more general terms, their theory predicts that better-informed investors will act as contrarians while less-informed investors will act as momentum traders. By trading at the expense of momentum traders, contrarians optimally allow momentum to persist, despite the existence of momentum traders.

The purpose of this paper is to test the theory proposed by Andrei and Cujean (hereafter AC) in the cryptocurrency market, where cryptocurrency is an alternative medium of exchange consisting of numerous decentralized crypto coin types including the most famous one - Bitcoin. Specifically, I show evidence that time-series momentum reflects the speed of asset-specific information percolation. In other words, both of the magnitudes of the short-term momentum effect and long-term reversal effect will increase with the meeting intensity measured by the Google trend index. These findings are robust across a number of look-back periods.

1.2 Cryptocurrency and Google trend index

In this section, I explain why cryptocurrency market is chosen to test AC's theory and why Google trend index is applied to proxy the frequency of meetings among investors (i.e., meeting intensity - the most important parameter in AC's theory).

Cryptocurrencies, instead of other traditional assets, are chosen to conduct the empirical test mainly because their prices appear to be at the mercy of nothing more than market sentiments (Weber, 2016; Dwyer, 2015). Recent literature (Abraham et al., 2018; Philippas, 2019) shows that the prices of various kinds of Cryp-

tocurrencies are largely influenced by online information (e.g., community views, Google Trends, number of tweets, and forum comments) as opposed to the price of the traditional financial assets, which are mostly determined by their fundamental value (e.g., cash flow and dividends). The nature of lacking fundamental values[2] and often being associated with speculative bubbles enables cryptocurrency a good candidate for AC's framework, in which the fundamental value of the asset is assumed to be unchanged during the trading periods and the market price of the asset is mainly determined by the word-of-mouth communication. Secondly, with low entry costs and plenty of publicly available information, cryptocurrency market has become a worldwide financial laboratory (Phillip et al., 2018; Weber, 2016; Yermack, 2013). Its innovative features like transparency, decentralization and no single overseeing authority have drawn much attention from media, government institutions and academic researches. A few examples are listed as follows: Chicago Mercantile Exchange (CME) began to trade futures on Bitcoin in 2017, J.P. Morgan launched its own cryptocurrency in 2019, and Facebook's cryptocurrency - Libra is projected to be released in 2020. In addition, there are currently over 5000 cryptocurrencies with a market capitalization of around 250 billion USD, and more than 20 thousand exchanges that are venues to a total daily trading volume that surpasses 130 billion USD (information obtained on 1 Feb, 2020 from CoinMarketCap.com). Thirdly, cryptocurrency market operates 24 hours a day without breaks, which provides better quality financial data (better in terms of no price jumps or discontinuity during weekends and holidays, which is very common in most of the traditional financial markets). This continuous trading feature also makes it possible for researchers to look more deeply into the variation of the momentum effect during a single day (e.g., higher meeting frequency during the daytime and lower during the nighttime), which will

[2] Cheah and Fry (2015) find empirical evidence that the fundamental price of one of the most famous cryptocurrency called Bitcoin is zero.

be explored more in Chapters 4 and 5. Last but not least, to the best of my knowledge, most of the previous studies within cryptocurrency markets employed daily data (Katsiampa et al., 2019; Koutmos, 2018; Fry and Cheah, 2016) and none of them have examined the relationship between meeting intensity and return serial correlation using high-frequency (minute level) data. This unique type of data, as it will prove lately, can provide cryptocurrency investors with better insights in terms of market behaviors and dynamics.

In markets like cryptocurrencies, I believe the relative frequency of Google search for a specific coin in a given time (known as Google trend index) is a well suited proxy for the meeting intensity parameter in AC's model. Since most of the cryptocurrencies are traded online and many coin developers and inventors share their opinions about the pros and cons of different coins in online forums (Garcia et al., 2014), many investors rely on the information on the internet to make decisions on buying and selling them (Panzarasa et al., 2009). One of the most common ways for them to obtain news and opinions online is through search engines, and Google is undisputedly the most popular search engine in the world. Instead of meeting offline, cryptocurrency investors are more used to searching and exchanging opinions online by using Google search engine (Hau and Kim, 2011). It means that when people search a keyword like "Bitcoin" for a specific day, it could be regarded as more frequent "online meetings" are happening regarding Bitcoin during that day. To sum up, as most interactions between cryptocurrency investors are online, and Google is the standard search engine for information, Google trend index should be a appropriate indicator for the "online meeting intensity".

1.3 Outline

The outline of this thesis is as follows. Chapter 2 describes the cryptocurrency and Google trend index data in details. Chapter 3 introduces AC's theory and derives the theoretical prediction of the relationship between return serial correlation and meeting intensity. Chapter 4 contains the main results on return serial correlation sorted by Google trend index and trading interval. Chapter 5 checks the robustness of the results. Chapter 6 concludes.

2 | Data and preliminaries

This chapter describes in details the various data sources I use. Moreover, I discuss the reason for choosing the specific data sources used.

2.1 Cryptocurrency returns

The data consist of minute-level closing prices from the 31st of January 2017 to the 27th of June 2019 for two pairs: Bitcoin to USD rate (hereafter BTC/USD) and Etherem[1] to USD rate (hereafter ETH/USD), which are the two largest cryptocurrency pairs in terms of market capitalization and trading volume during the sample period. Furthermore, Bitcoin was the first cryptocurrency ever created in the world and was released in 2009. As a result, these two instruments are currently among the most liquid cryptocurrency pairs in the world. I choose these liquid instruments in order to avoid the stale price problem in the high-frequency data. I use high-frequency data to test AC's theory because algorithmic trading, a methods of trading using automated preprogrammed instruction that heavily depends on the speed and a large amount of fine-grained financial data, nowadays dominates the trading scene in the financial markets, particularly cryptocurrency market (Fang et al., 2020). Moreover, high-frequency data is more informative compared to the daily data and may yield a more clear picture regarding portfolio analysis and risk management. From the point of the efficiency market hypothesis, high-frequency data has the edge over the traditional daily data in the empirical

[1] Ethereum is a blockchain platform and Ether (ETH) is the currency that is transacted through the platform. Users receive the Ether currency in return for contributing their computing power in validating transactions and aiding in the development of the platform. The Ether currency is also used to pay for "specific actions on the Ethereum network". In the remainder of this thesis, I refer to the Ether currency as Ethereum, as this is the standard procedure in other literature and online cryptocurrency exchanges.

study of informational efficiency, which will be discussed in more detail in Section 4.1 after I introduce the actual model.

Return series for each instrument is calculated using minute return: $r_t = (p_t - p_{t-1})/p_{t-1}$ where r_t is the return series, p_t is the price series, and p_{t-1} is the one lag price series. I also compound the minute returns to a cumulative return from which returns at any horizon can be easily calculated. Notice here the analysis is limited to the data that were available to download at the time of writing. Data sources detail can be found in Appendix A.

Table 2.1 presents the summary statistics of the prices and returns on the two pairs. The columns report the time series mean, standard deviation, minimum and maximum in the frequency of every minute. As Table 2.1 highlights, the sample mean returns for both pairs are very close to zero and they both have abnormal fat tail return distributions. Not surprisingly, the return volatility for ETH/USD is much larger than for BTC/USD due to ETH's smaller capitalization during the sample periods. However, the BTC/USD minute return varies over a broader range (-17.5% to 12.1%) than ETH/USD (-10% to 5.9%), perhaps due to a series of Bitcoin specific events happened in the sample period (e.g, in 2017, Japan declared Bitcoin a legal currency and US Commodities Futures Trading Commision approved Bitcoin futures, in 2018, South Korean government officials were caught insider trading and manipulating the Bitcoin market). These features of high price volatility and heavy tails make cryptocurrency an inferior way to store real value, which makes it different from traditional currencies (Gkillas and Katsiampa, 2018). Nevertheless, cryptocurrencies can be viewed as a new kind of tradable speculative asset that heavily rely on public sentiments, which makes it well suited to evaluate AC's model.

Table 2.2 shows the stationary properties of the price time series for the two pairs.

Figure 2.1: Bitcoin and Etherem every minute prices from the 31st of January 2017 to the 27th of June 2019

As expected, standard Augmented Dickey-Fuller (ADF) tests strongly indicate that the price series have a unit root, whereas the first differences of the prices (i.e., the returns) are stationary. As a robustness check, further tests based on different assumptions concerning the deterministic component (none/constant/trend) have been conducted and yield the same conclusions as before.

Figure 2.1 shows every minute prices for Bitcoin (red curve) and Etherem (blue curve) from the 31st of January 2017 to the 27th of June 2019. In 2017 the USD prices of the two cryptocurrencies showed an explosive rise, while the year 2018 and the first half of the year 2019 were marked by bearish markets. For the overall sample, BTC/USD prices ranged from $ 744 to $ 19,891 and ETH/USD prices ranged from $ 8 to $ 1,423. Arguably, the two cryptocurrencies compete with each other and share the same price formation determinants. Therefore their prices should be closely related, as can be seen in the figure.

Table 2.1: Summary statistics of the price and returns on BTC and ETH from the 31st of January 2017 to the 27th of June 2019

Statistic	Mean	St. Dev.	Min	Max
Price_BTC	5,765	3,443	744	19,891
Return_BTC	0.00000	0.001	−0.175	0.121
Price_ETH	345.723	258.990	8.044	1,423.400
Return_ETH	0.00001	0.002	−0.100	0.059

Table 2.2: Augmented Dickey-Fuller test statistics of the price and returns on BTC and ETH (from the 31st of January 2017 to the 27th of June 2019)

Statistic	Prices	Returns
BTC	-1.713 (0.702)	-112.38 (0.000)
ETH	-2.05 (0.559)	-108.9 (0.000)

Note: The ADF test regression equation incorporates a constant and a linear trend. The t-statistic and p-value (in parentheses) are reported in the table.

2.2 Google trend index

In the modern world, nearly every aspect of daily life involves the internet. Most of the people are getting used to navigating the virtual world via search engines. Amongst these search engines, Google is one of the most popular choices. In fact, Google has 1.5 billion users worldwide, and processes an average of around 40,000 search queries per second (Abraham et al., 2018). It implies that Google search data can shed some light on who is interested in what, and how trends in some given topics are evolved.

Google makes this type of data available through "Google trend index". It is a search trend feature that can provide users with how frequently a specific search key term is entered into Google's search engine over a given period of time. Google trend index can be used for comparative keyword research and to discover

event-triggered spikes in keyword search volume. It can also provide keyword-related data, including search volume index at any given time. This function provides a good proxy for the "meeting intensity" parameter in the AC's model.

Since Google uses a search volume index (SVI) to calculate the Google trend index, and the data queried for a period of longer than 90 days are automatically aggregated at a weekly level, some adjustment has been made to compare the SVIs at a daily level. Adjustment method detail can be found in appendix A.

Table 2.3: Summary statistics of adjusted Google Trends (AGT) using "bitcoin" and "ethereum" as keywords

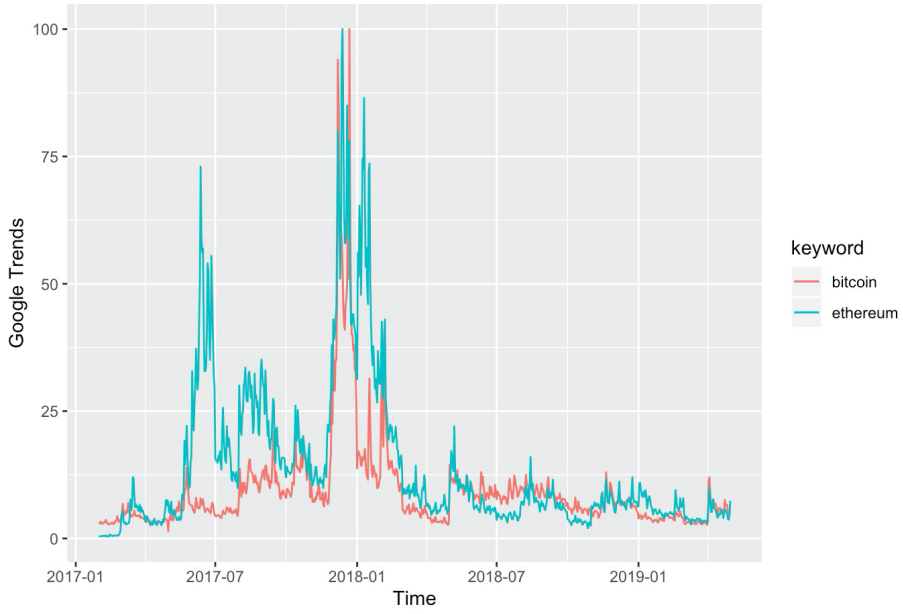
Google Trends	Mean	St. Dev.	Min	Max
bitcoin	9.214	9.628	1.440	100.000
ethereum	13.942	15.360	0.360	100.000

Note: Since Google is using a search volume index (SVI) to calculate the Google Trends data, and the data queried for a period of longer than 90 days are automatically aggregated at a weekly level, some adjustment has been made to compare the SVIs at a daily level and the adjusted Google Trend (AGT) is reported in the table. Adjustment method detail can be found in appendix A.

In order to avoid ambiguity, currency's abbreviation (i.e., "BTC" and "ETH") were not used in the keyword search (e.g., ETH could ambiguously mean ETH Zurich university). Instead, I collect the Google trend index data using the keyword "bitcoin" and "ethereum" (case insensitive) separately.

Table 2.3 presents summary statistics of the daily adjusted Google Trends (AGT) using the keyword "bitcoin" and "ethereum" from the 31st of January 2017 to the 27th of June 2019. The columns report the time series mean, standard deviation, minimum and maximum in the frequency of everyday. By construction, both of the "bitcoin" and "ethereum" AGT are scaled between 0 and 100. It is noteworthy that AGT time series with different keywords are not comparable due

Figure 2.2: Adjusted Google Trends (AGT) values for Bitcoin and Ethereum daily data from the 31st of January 2017 to the 27th of June 2019



to the Google internal scale normalization. Therefore comparisons between the "bitcoin" AGT and "ethereum" AGT time series in Table 2.3 are meaningless. However, the time variations of the AGT series can still be safely used to explain the variations in return series respectively in the following regression analysis.

Figure 2.2 shows the daily AGT time series using the keyword "bitcoin" (red curve) and "ethereum" (blue curve) from the time periods. It is clear to see the 2 AGT time series are closely related, suggesting that people who are interested in Bitcoin activities also want to know more about other cryptocurrencies.

3 | Andrei and Cujean's Theory

This chapter provides a brief introduction to the Andrei and Cujean (2017)'s joint theory of time-series momentum and reversal so the readers can attain a better theoretical motivation for the hypothesis tests in Chapter 4.

Section 3.1 defines the objective function of the agent in the financial market and describes the basic structure of the economy. Section 3.2 describes how information percolates around agents. Section 3.3 solves the model and yields the equilibria for the economy, which consist of the financial asset's price and investors' demand for every period. Finally, Section 3.4 derives the most important prediction of the theory based on the models and propositions developed in the first three sections: the theoretical relationship between the meeting intensity parameter and the return serial correlation. The relationship is proved to be non-linear in meeting intensities and varies across different lags.

3.1 Economy

Assuming investors have exponential utility with a common coefficient of absolute risk aversion $1/\tau$, where τ denotes investors' risk tolerance - the smaller τ means less risk tolerance, therefore more risk-averse. Investors can choose to trade a single asset, whose payoff $\tilde{U}$ is realized at time $t = T$. Trading takes place at time $t = 0, 1, ..., T-1$ and market price $\tilde{P}_t$ is endogenously determined by the trading among investors. Each investor i is endowed at time $t = 0$ with a quantity of the risky asset represented by X^i. At each trading date, investor i chooses a position in the risky asset, $\tilde{D}_t^i$, in order to maximize her expected utility of terminal wealth, denoted by $\tilde{W}_T^i$:

$$\max_{\tilde{D}_t^i} E[-e^{-\frac{1}{\tau}\tilde{W}_T^i}|\mathcal{F}_t^i] \quad s.t.$$

$$\tilde{W}_T^i = X^i\tilde{P}_0 + \sum_{t=0}^{T-2}\tilde{D}_t^i(\tilde{P}_{t+1} - \tilde{P}_t) + \tilde{D}_{T-1}^i(\tilde{U} - \tilde{P}_{T-1}) \quad (3.1)$$

The terminal wealth $\tilde{W}_T^i$ can be decomposed into two parts: the value of investor i's endowed asset at time 0 ($X^i\tilde{P}_0$); and the profit or loss that investor i makes via trading from time 0 to time T, which is determined by the position in the risky asset at every trading period and the market price movement (the last period is slightly different as the underlying payoff $\tilde{U}$ is realized).

As every investor seeks to solve the optimization problem and make trading decisions based on his own information, it is important to understand what this information set contains. Andrei and Cujean assume the information set of investor i at time t, $\mathcal{F}_t^i$, contains two parts - the public and private signals. The public signal is represented by the endogenously determined market prices in equilibrium, which is a common information diffusion mechanism in a standard rational-expectation framework (Grossman and Stiglitz, 1980). The second part is the private signals that are collected through information percolation. This information percolation structure is a word-of-mouth communication mechanism (Duffie and Manso, 2007) that can enforce information flowing at an increasing rate and can generate heterogeneous trading strategies among agents. Section 3.2 will provide more details about how the two channels (public and private) together form the agent's information set over time.

Next, the noisy liquidity traders are introduced. At time $t = 0$, the aggregate per capita supply of the risky asset $\tilde{X}_0 = \int_0^1 X^i di$ is assumed normally and in-

dependently distributed as zero mean and precision Φ, where the precision of a random variable is referred to the inverse of its variance. New liquidity traders enter the market buying/selling the risky asset randomly in trading sessions $t = 0, 1, ..., T-1$. The incremental net supply of liquidity traders represented by $\tilde{X}_t$, is also assumed normally distributed as zero mean and precision Φ. Noisy liquidity traders are needed because the noisy supply prevents asset prices from fully revealing the final payoff $\tilde{U}$[1].

3.2 Information percolation

Since the risky security payoff $\tilde{U}$ is unobservable and assumed to follow a normal distribution with zero mean and precision H, investors need information from the public market prices and the private meetings to enhance the precision of their guesses about the risky security payoff. It is assumed that immediately prior to trading session $t = 0$, each investor i obtains a private signal about the asset payoff $\tilde{z}^i = \tilde{U} + \tilde{\epsilon}^i$ where $\tilde{\epsilon}^i$ is distributed normally and independently of $\tilde{U}$, has zero mean and precision S.

The meeting process is introduced in order to add a channel of information acquisition for the agents in the economy. Andrei and Cujean model word-of-mouth communication through the information percolation theory (Duffie and Manso, 2007), whereby agents exchange information in random, bilateral private meetings. The meeting process is set up in a way so that information can flow at an increasing rate and information precision could become heterogeneous across agents. Specifically, from date $t = 0$ onward, agents meet each other randomly

[1] A random walk specification for the noisy supply is adopted (i.e., the total supply at time $t$is $\Sigma_{j=0}^{t}\tilde{X}_j$). Under this specification and in the absence of additional private information at dates $t \geq 1$, prices are martingales.

and share their information. Meetings take place continuously at Poisson arrival times with meeting intensity λ - the only parameter that is added to the standard equilibrium model (Grossman and Stiglitz, 1980). λ can be interpreted as the meeting frequency, the larger the value of λ, the more frequent the meetings occur, hence information will percolate faster. As will be demonstrated in Section 3.4, the magnitude of meeting intensity can not only determine whether the asset returns exhibit momentum or reversal, but also the magnitude of these phenomena.

Between trading dates $t-1$ and t, each agent i collects a number $\omega_t^i \in \mathbb{N}$ of signals. The cross-sectional distribution of the number of additional signals from trading dates $t-1$ and t is denoted as $\pi_t(n)$ where n is the number of additional signals. For example, since agents are initially endowed with a single signal, the initial distribution of signals has 100% probability mass at $n=1$. Therefore it is easy to show that $\pi_{t=0}(n)=1$ for $n=1$ (since at $t=0$ every agent has exact one signal) and $\pi_{t=0}(n)=0$ for all $n \in \mathbb{N}$ and $n \neq 1$. Moreover, the cross-sectional average of the number of additional signal at time t is $\Omega_t \equiv \Sigma_{n\in\mathbb{N}}\pi_t(n) * n$. This statistic is important because it describes the overall level of investors' information regarding the risky security payoff, which is later proved a determinant of market prices. Furthermore, Andrei and Cujean prove that, under the Poisson process assumption, the cross-section average has such relationship with time and the meeting intensity parameter: $\Omega_t = e^{(t-1)\lambda}(e^{\lambda}-1)$. The result shows two things: the larger the meeting intensity, the larger the average of the additional signals across different agents as information percolates faster. Secondly, as time goes by (t increases), the average of the additional signals also increases because on average agents will hold more signals day by day through sharing private signals in every meeting.

To represent the information in a more precise way, it is convenient to relate the

number of signal (ω_t^i) to the precision of agent's guess about the risky security payoff $\tilde{U}$. Since at each date t, agent i receives ω_t^i number of new signals with the same precision S. From Gaussian theory, these signals are equivalent to a single signal with precision $S\omega_t^i$. The integrated signal is denoted by $\tilde{Z}_t^i$ - the aggregate additional signal agent i obtain in period t. It can be shown that $\tilde{Z}_t^i = \tilde{U} + \bar{\varepsilon}_t^i$, where $\bar{\varepsilon}_t^i \equiv (\frac{1}{\omega_t^i}\sum_{j=1}^{\omega_t^i}\tilde{\varepsilon}^j) \sim N(0, \frac{1}{S\omega_t^i})$, representing the cross-section aggregate signals from different investors that have met with agent i. The precision of $\tilde{Z}_t^i$ is then the inverse of the variance therefore $S\omega_t^i$. It implies that the more numbers of signals an agent receives in a period, the more precise the combined single signal ($\tilde{Z}_t^i$) will be at that specific period. Taken together, the conditional precision of the agent i about the final payoff $\tilde{U}$ is defined as $K_t^i \equiv Var^{-1}[\tilde{U}|\mathcal{F}_t^i]$, where $\mathcal{F}_t^i$ is the total information set of investor i at time t. The information set contains again two parts: the cumulative private signals collected through the word-of-mouth communication during all of the past time periods, mathematically denoted as $\{\tilde{Z}_j^i\}_{j=0}^t$. And the cumulative public signals represented by the historical trading prices, denoted as $\{\tilde{P}_j\}_{j=0}^t$. Just like the cross-sectional average of the number of additional signal at time t is defined as $\Omega_t \equiv \Sigma_{n\in\mathbb{N}}\pi_t(n) * n$, the cross-sectional average of conditional precisions over the entire population of agents is defined as $K_t \equiv \Sigma_{n\in\mathbb{N}}\pi_t(n) * K_t^i(n)$. This statistic is important because it describes the overall precision of investors' information about the risky security payoff, which is proved later a determinant of market prices.

To sum up, via meetings, individual investors exchange their private signals and acquire a better guess about the security final payoff. More importantly, their information about the payoff over time can be summarized mathematically by a single metric K_t^i - the conditional precision of the agent i about the final payoff $\tilde{U}$. In addition, the overall precision of investors' information about the risky security payoff can be summarized by another metric K_t - the cross-sectional average of

conditional precisions over the entire population of agents.

3.3 Equilibrium

The equilibrium price and demand of the above economy under the word-of-mouth communication is defined in Theorem 1 (Equilibrium)[2] below:

Theorem 1: There exists a partially revealing rational-expectations equilibrium in the economy in which the price of the risky asset, $\tilde{P}_t$ for $t = 0, 1, ..., T-1$, is given by:

$$\tilde{P}_t = \frac{K_t - H}{K_t}\tilde{U} - \sum_{j=0}^{t} \frac{1 + \tau^2 S\Omega_j \Phi}{\tau K_t}\tilde{X}_j. \tag{3.2}$$

The individual asset demands, $\tilde{D}_t^i$, are given by

$$\tilde{D}_t^i = \tau K_t^i (E[\tilde{U}|\mathcal{F}_t^i] - \tilde{P}_t). \tag{3.3}$$

The individual and average market precision are given by

[2]The solution method for finding a linear, partially revealing rational-expectations equilibrium is standard and is showed in Appendix A.2 (Andrei and Cujean, 2017).

$$K_t^i = H + \sum_{j=0}^{t} S\omega_j^i + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi,$$
$$K_t = H + \sum_{j=0}^{t} S\Omega_j + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi.$$

The precision of the econometrician is given by

$$K_t^c = H + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi.$$

Theorem 1 describes the risky asset prices at each date in a noisy rational-expectations equilibrium with in formation percolation.

The asset price in Eq. (3.2) is a linear function of the final payoff and supply shocks. The higher the final payoff, the higher the equilibrium price. The coefficient $(K_t - H)/K_t$ will converge to 1 as time goes by, implying that ignoring the net supply shock, the equilibrium price will become more and more precise (i.e., reflect the true value of the final payoff) over time. The second term indicates that the more noise traders sell (i.e., the net supply increases) up till to t, the lower the equilibrium price at t.

The second part of the theorem (i.e., Eq. (3.3)) shows how the demand of individual i at t is determined. It depends on i's risk tolerance (τ), his individual precision (K_t^i), and the difference between the expected final payoff based on his information set and the actual market price. Intuitively, if an investor believes an

asset is overvalued (i.e., $\tilde{P}_t > E[\tilde{U}|\mathcal{F}_t^i]$), he will sell it. The more confident he is (higher K_t^i), the more risk he is willing to take for the transaction (higher τ), therefore he will sell more of it. If an investor believes an asset is undervalued, he will buy it. The more confident he is (higher K_t^i), the more risk he is willing to bear for the transaction (higher τ), therefore he will buy more of it.

For the precision part, by definition, we know $K_t^i \equiv Var^{-1}[\tilde{U}|\mathcal{F}_t^i]$. As we have discussed in the end of the section 3.2, the information set contains the cumulative private signals and the cumulative public signals, we have $K_t^i = Var^{-1}[\tilde{U}|, \{\tilde{P}_j, \tilde{Z}_j^i\}_{j=0}^t]$. Therefore the precision about the payoff of agent i at time t conditional on the total information set depends on three things: H - the precision of the payoff; $\sum_{j=0}^t S\omega_j^i$ - the precision of the cumulative private information of agent i; $\sum_{j=0}^t \tau^2 S^2 \Omega_j^2 \Phi$ - the precision of the cumulative public information. The last term could be easily derived using the variance of the the normalized price signals. The intuition is that the smaller variance of the payoff distribution, the more precise the market price is able to reflect the final payoff, hence the individual i will acquire a higher precision in the end[3]. Once we understand the individual precision, we could extend the intuitions to the market precision (i.e., change the individual additional number of signal at t to the average across agents) and the econometrician precision (i.e., mute the private signals channel).

3.4 Meeting intensity and return serial correlation

After solving the equilibrium market price and individual's demand for the asset, we can now show the relationship between the meeting intensity parameter and

[3]For more details please refer to page 634, Appendix A.2 (Andrei and Cujean, 2017).

the return serial correlation in proposition 1[4]:

Proposition 1: Conditional on past returns, expected future returns satisfy

$$\mathbb{E}[\tilde{R}_{t+1}|\mathcal{F}_t^r] = \sum_{l=1}^{t+1} \frac{K_{t+1} - K_t}{K_{t+1}K_t^c} m_{t-l}(\tilde{R}_{t-l+1})$$

where $K_t = H + \sum_{j=0}^{t} S\Omega_j + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi$,

$\tilde{R}_t = \tilde{P}_t - \tilde{P}_{t-1}$ (returns between t and $t-1$),

$K_t^c = H + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi$,

$\mathcal{F}_t^r = \{\tilde{P}_{t-l+1} - \tilde{P}_{t-l} : 1 \leq l \leq t+1\}$,

$\Omega_t = e^{(t-1)\lambda}(e^{\lambda} - 1)$ and

$$m_{t-l} \equiv \{\underbrace{\sum_{k=0}^{t-l} S\Omega_k}_{momentum}\} - \{\underbrace{\frac{S\Omega_{t-l+1}}{(K_{t-l+1} - K_{t-l})/K_{t-l}}}_{reversal}\} \tag{3.4}$$

It is essential to understand the sign of the serial correlation coefficient because the sign will determine whether returns exhibit momentum or reversal. Notice that the average precision is always increasing over time, leading to $K_{t+1} - K_t > 0$. Knowing $K_{t+1} > K_t$, the sign of the coefficient for $\tilde{R}_{t-l+1}$ will be solely determined by m_{t-l}. If m_{t-l} is positive, the return for period t and lag l will exhibit momentum. Otherwise the return will exhibit reversal. Therefore $m_{t-l} > 0$ is called the "momentum condition" for period t and lag l. In other words, the sum of the reversal effect and the momentum effect in Eq. (3.4) determines the sign of

[4]The proof is in Appendix B - B4 (Andrei and Cujean, 2017)

the serial correlation of stock returns at the lth lag.

As can be seen from Eq. (3.4), the momentum effect strengthens as private signals "accumulate". For the reversal component, the magnitude of the effect strengthens with the accumulation of private signals (numerator), but weakens as average market precision increases (denominator). For momentum to arise, the increase in average market precision must be sufficiently large to restore the balance in favor of the momentum effect. To achieve a better intuition of the momentum condition, this momentum condition is redefined in terms of the precision elasticity. Appendix B shows an equivalence between $m_t > 0$ and $\epsilon_t > 1$ where

$$\epsilon_t \equiv \frac{(K_{t+1} - K_t)/K_t}{(\sum_{k=0}^{t+1} S\Omega_t - \sum_{k=0}^{t} S\Omega_t)/\sum_{k=0}^{t} S\Omega_t}.$$

ϵ_t is defined as the t period precision elasticity - a general concept that characterizes the "shape of learning" in a rational-expectations model. It measures how average market precision responds to a change in the average precision of private information. When the meeting intensity is relatively large, the private information flows at a steep increasing rate[5]. Hence the average market precision increases faster than the increase in the average precision of private information[6], therefore the elasticity is greater than one and the return exhibits momentum.

Substituting $K_t = H + \sum_{j=0}^{t} S\Omega_j + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi$ and $\Omega_t = e^{(t-1)\lambda}(e^\lambda - 1)$ into the momentum condition and keeping the other parameters H, S, Φ, τ unchanged, we can derive the values of λ that satisfy the momentum condition. Andrei and Cujean (2017) show that there exists a unique threshold of the meeting intensity above which stock returns always exhibits momentum. They call this threshold

[5] We can regard the denominator as the speed of the average precision of the private information. The second derivative of $\Omega_t = e^{(t-1)\lambda}(e^\lambda - 1)$ w.r.t t is $e^{(t-1)\lambda}\lambda^2(e^\lambda - 1)$: the acceleration speed of the flow of private information is at least quadratic in λ.

[6] Since $K_t = H + \sum_{j=0}^{t} S\Omega_j + \sum_{j=0}^{t} \tau^2 S^2 \Omega_j^2 \Phi$ - the price tomorrow incorporates increasingly precise information about the fundamental. Not only does momentum require the price-learning channel for average market precision to increase faster than the flow of private information, it also requires the flow of private information to have the "right dynamics." (specific values of λ).

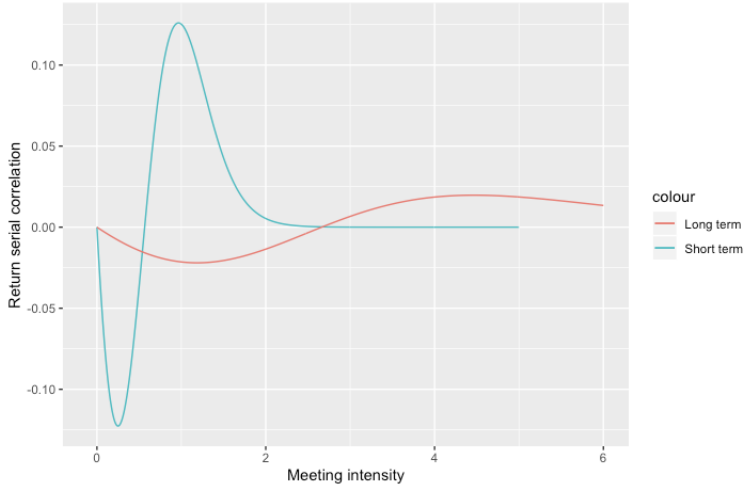
the momentum threshold and denote it as λ^*.

Not only can we figure out how meeting intensity could influence the sign of return serial correlation, but we are also able to derive the exact relationship between the magnitude of the meeting intensity parameter and the return serial correlation in different lags. Appendix B shows the return serial correlation of lag l is a function of parameters $\lambda, l, H, S, \Phi, \tau$ where again λ is the Poisson distribution parameter (meeting intensity), H is the precision of the security payoff, S is the precision of the signal distribution, Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock), and τ is the investor risk tolerance.

As for a fixed lag l, the return serial correlation is nonlinear in λ, by using the same calibrations as in Andrei and Cujean (2017) for the hyper-parameters ($S = H = \Phi = 1, \tau = 1/3$) we can obtain the non-linear relationship between the meeting intensity parameter and the return serial correlation, which is plotted in Figure 3.1. Serial correlation is computed for two different lags: the blue line corresponds to lag $l = 1$, representing the short-term correlation; and the red line corresponds to lag $l = 60$, representing the long-term correlation.

From 3.1, we first notice that without information percolation ($\lambda = 0$), returns are unpredictable. With information percolation ($\lambda > 0$), stock returns exhibit reversals when the meeting intensity is below the momentum threshold (λ^*). As information percolation intensifies and the meeting intensity rises above the threshold, stock returns become positively autocorrelated. An increase in the meeting intensity beyond this threshold weakens the reversal effect and strengthens the momentum effect, thereby creating a stronger relation between momentum and the meeting intensity. As the meeting intensity increases further, not only does the reversal effect die out, but the momentum effect also disappears, resulting in a negative relationship between momentum and the meeting intensity. Finally, when

Figure 3.1: Relationship between return serial correlation and meeting intensity



Information percolation and serial correlation in returns. Serial correlation of returns as a function of the meeting intensity λ. Serial correlation is computed for two different lags: the blue line corresponds to lag $l = 1$, and the red line to lag $l = 60$. The calibration used is $S = H = \Phi = 1$, and $\tau = 1/3$.

meeting intensity becomes infinite, returns become serially uncorrelated. All together it results in a hump-shaped momentum pattern when the meeting intensity exceeds the momentum threshold.

Instead of looking at a single curve over meeting intensity, now we focus more on comparing the two curves (red and blue) with different lags. We observe that the momentum threshold increases with the lags since the long-term (red) curve has a larger value of momentum threshold than the short-term (blue) curve. Moreover, both of the magnitudes of the peak and trough for the short-term curve are higher than the long-term curve.

4 | Hypothesis testing: momentum effects, cut different ways

This chapter displays the main hypothesis testing results in cryptocurrency market based on the Andrei and Cujean (2017) 's theoretical predictions. In the most general terms, I list the two central hypotheses of this thesis based on the three assumptions below:

The three assumptions are:

Assumption 1: During the sample period, the theoretical meeting intensity parameter λ can be proxied by the Google trend index.

Assumption 2: During the sample period, the short-term effect can be represented by the one minute lag and the long-term effect can be represented by the one hour lag (60 minutes).

Assumption 3: During the sample period, H (precision of the security payoff), S (precision of the signal distribution), Φ (precision of the aggregate per capita supply of the risky asset), and τ (investor risk tolerance) do not change.

The two hypotheses are:

Hypothesis 1: Given assumptions 1 - 3, asset returns should display more positively autocorrelated returns when meeting intensity is high at a short-term horizon.

Hypothesis 2: Given assumptions 1 - 3, returns should display more negatively autocorrelated returns when meeting intensity is high at a long-term horizon.

The first assumption is considered to be reasonable given the arguments mentioned in Section 2.2. The second assumption will be discussed more in Section 4.1. The third assumption is more difficult to validate, but on average should hold true in a high-frequency data setting. That being said, extra cautions should be taken in order to interpret the testing results.

4.1 Cuts on meeting intensity

This section starts by quantifying the time-series momentum effect of returns across different time horizons. Following the procedure in Andrei and Cujean (2017), I regress the minute return r_t^s for cryptocurrency pair s in time t on its return lagged h periods:

$$r_t^s = \alpha + \beta_{s,h} r_{t-h}^s + \epsilon_t^s \tag{4.1}$$

Where $\beta_{s,h}$ is the return serial correlation for pair s and lag h, and ϵ_t^s is the statistical noise. The regressions are run using lags of $h = 1$ minute for the short-term effect and $h = 1$ hour (60 minutes) for the long-term effect. Notice that the definition of the short-term and long-term periods in our data setting (high-frequency data) is quite different from the traditional financial data setting, which normally uses the relatively low frequency data (mostly daily data). For example, Moskowitz et al. (2012) define the short-term frequency from one to six months and the long-term frequency over twelve months. On the contrary, I am using a much higher sampling frequency in the thesis because situations change a lot

nowadays. Due to the increasing popularity of ultra high-frequency algorithmic tradings in cryptocurrency market, trading frequency has been greatly levelled up. Therefore the definition of "short-term" and "long-term" should also be adjusted (Makarov and Schoar, 2020; Lewis, 2014)[1]. In the recent high-frequency trading literature, Aslan and Sensoy (2019) have examined the efficiency in cryptocurrency market for different time intervals using minute data by applying Hurst exponents tests. They have found that the high-frequency traders who use algorithmic trading strategies can generate abnormal profits in the cryptocurrencies by trading on 1 minute or 1 hour basis, which implies substantial information inefficiency in these two frequencies. The information inefficiency in these two frequencies would then be able to provide strong power to test the sign of $\beta_{s,h}$ in Eq. (4.1). In addition, Chapter 5 constructs more sensitivity tests in different sampling frequencies in order to check the robustness of the results.

Next I present the results of the estimates for $\beta_{s,h}$ using different subsamples, which closely follow the structure of Hong et al. (2000) in many respects. I use an independent sort to generate 4 subsamples by the meeting intensity proxy (AGT) - the low subsample which includes observations below the first quartile of the meeting intensity distribution, the mid-low subsample which includes the first quartile to the second quartile, the mid-high subsample which includes the second to third quartile and the high subsample which includes observations above the third quartile. Table 4.1 shows the results of the estimates of $\beta_{s,h}$ using different lags in the above 4 subsamples.

The first observation from Table 4.1 is that the estimating results confirm the typical short-term momentum and long-term reversal phenomenon in the full sample, as can be seen from the first row. The one-minute lag return shows the positive predicting power for the future return while the one-hour lag return shows the

[1]For a more cryptocurrency specific discussion, see Daian et al. (2019)

Table 4.1: Momentum effect using Bitcoin returns and sorting by meeting intensity

	short-term	long-term	short - long
All	0.041***	-0.007***	0.048***
	(0.001)	(0.001)	(0.000)
Low MI	0.009**	-0.002	0.011***
	(0.002)	(0.002)	(0.000)
Mid-low MI	0.032***	0.000	0.032***
	(0.002)	(0.002)	(0.000)
Mid-high MI	0.062***	0.002	0.06***
	(0.002)	(0.002)	(0.000)
High MI	0.066***	-0.011***	0.077***
	(0.002)	(0.002)	(0.000)
High - Low	0.057***	-0.009***	0.066***
	(0.000)	(0.000)	(0.000)

Note: This table lists the results of the estimates of $\beta_{s,h}$ in model 4.1 for different return lags. The first column lists the estimates of $\beta_{s,h}$ using lags of $h = 1$ minute for the short-term momentum effect. The second column lists the estimates of $\beta_{s,h}$ using lags of $h = 1$ hour for the long-term momentum effect. The last column shows the quartile of the meeting intensity (MI) distribution. The rows report the estimates using an independent sort on the meeting intensity (proxied by AGT). The low subsample (Low MI) includes observations from 0 to the first quartile (MI_{Q_1} = 4.62) of the meeting intensity distribution, the mid-low subsample includes the first quartile (4.62) to the second quartile (7), the mid-high subsample includes the second (7) to third (10.4) quartile and the high subsample includes observations above the third quartile (10.4). The third column and the last row are calculated using the two-sample t-test from the mean, standard deviation and sample size of the individual estimates. The number of observations for each subsample is 288,266. Standard errors are in parentheses. *** p < 0.001; ** p < 0.01; * p < 0.1

negative predicting power. Both of the p-values are less than 0.1%, implying the testing results are highly significant.

More interestingly, the first and second column shows how the short and long-term momentum magnitude changes when the meeting intensity changes. The first column shows there is a pronounced positive relationship between the meeting intensity and the magnitude of the momentum effect at a short-term horizon. Looking more deeply into every row of the first column, we can see that when meeting intensity is low with AGT less than the first quartile (4.62), the short-term time-series momentum effect is actually very limited with a value of 0.009. However, the magnitude of the momentum in the mid-low subsample surges to 0.032, which is 3.6 times of the magnitude of the low subsample. The momentum effect continues to almost double its magnitude from the mid-low subsample to the mid-high subsample and gradually increase to its peak (0.066) at the high subsample. The last row shows the magnitude of the momentum effect with the large meeting intensity is significantly higher than the magnitude with the low meeting intensity. The difference of 0.057 between the two subsamples is highly statistically significant, with a p-value smaller than 0.1%. Therefore the Hypothesis 1 that asset returns should display more positively autocorrelated returns when meeting intensity is high at a short-term horizon is validated. This positive relationship is again, in line with the Andrei and Cujean (2017)'s theory which predicts that with a small lag (short-term), an increase in the meeting intensity beyond the momentum threshold weakens the reversal effect and strengthens the momentum effect, thereby creating an increasing relation between momentum and the meeting intensity.

The second column shows a roughly negative relationship between the long-term momentum and the meeting intensity. In other words, when meeting intensity increases, the magnitude of the reversal effect becomes more salient when one-hour

(long-term) returns are used. The momentum measure is -0.002 in the low meeting intensity subsample and it decreases to -0.011 in the high meeting intensity subsample. The difference of -0.009 between the two subsamples in this regard is highly statistically significant, with a p-value smaller than 0.1%. Therefore the Hypothesis 2 that asset returns should display more negatively autocorrelated returns when meeting intensity is high at a long-term horizon is also validated. This negative relationship is consistent with the Andrei and Cujean (2017)'s theory which predicts that with a long lag (long-term), an increase in the meeting intensity below the reversal turning point (therefore below the momentum threshold)[2] weakens the momentum effect and strengthens the reversal effect, thereby creating a decreasing relation between momentum and the meeting intensity. The seemingly contradicted conclusions between the positive relationship in the previous paragraph and the negative relationship in this paragraph are actually reasonable because the momentum threshold is also changed – it increases with the lag.

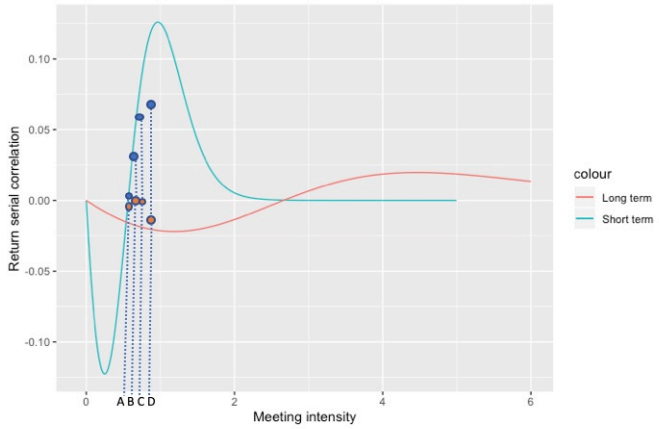
The third column shows that for any subsample, the short-term momentum is higher than the long-term momentum effect at 0.1% level of significance, which is again, consistent with Andrei and Cujean (2017)'s theoretical prediction. Assuming the investors' risk tolerance is fixed, the momentum effect is a function of the meeting intensity and look-back periods (lags). It can be shown that when meeting intensity is fixed, the momentum effect has a decaying "term-structure" whereby returns exhibit momentum for short look-back periods and reversal over longer look-back periods[3].

To better illustrate the different point estimates in Table 4.1, I plot the serial correlation of returns as a function of the meeting intensity for different lags using

[2]The reversal turning point is the stationary point in the phrase where stock returns exhibit reversal. Before this point, returns serial correlation decreases with meeting intensity.

[3]See the Model Predictions and Empirical Evidence section in Andrei and Cujean (2017) for more details

Figure 4.1: Simulated relationship between meeting intensity and serial correlation in returns with point estimates from Table 4.1



Note: The blue curve is drawn by the simulated points from the AC's model using lag $h = 1$ (short-term). The red curve is drawn by the simulated points using lag $h = 60$ (long-term). The calibration used in the figure is the same as in Andrei and Cujean (2017): $H = S = \Phi = 1$ and $\tau = 1/3$ where H is the precision of the security payoff, S is the precision of the signal distribution, Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock), τ is the investor risk tolerance. The four levels of meeting intensity A, B, C and D stand for the respective rows in Table 4.1: Low MI, Mid-low MI, Mid-high MI and High MI.

the standard calibration as in Andrei and Cujean (2017)[4]. In Figure 4.1, the blue curve is drawn by the simulated points from the derivations of AC's model using lag $h = 1$ (short-term). The red curve is drawn by the simulated points using lag $h = 60$ (long-term). In addition, I have plotted the four levels of meeting intensities A, B, C and D, which stand for the respective rows in Table 4.1: Low MI, Mid-low MI, Mid-high MI and High MI. It is important to know that the x-axis in the graph is using the actual Poisson arrival rate which assumes meetings take place continuously at Poisson arrival times with a single parameter - meeting intensity. This is the exact same setting as the word-of-mouth communication model[5] in Andrei and Cujean (2017). Caution should be taken when comparing the cardinal value of the Poisson arrival rate (actual meeting intensity parameter in the model) and the ordinal rank from the Google trend index. That being said, we can see a close match between the empirical estimates (dots) and the theoretical prediction (simulated curves), especially for the estimates that are statistically significant (i.e., all blue points and red point D).

In conclusion, Table 4.1 points to a consistent positive effect of meeting intensity proxy (AGT) on an asset's serial correlation at a short-term horizon and negative effect at a long-term horizon.

[4]The calibration used is $H = S = \Phi = 1$ and $\tau = 1/3$ where H is the precision of the security payoff, S is the precision of the signal distribution, Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock), τ is the investor risk tolerance in the Andrei and Cujean (2017)'s rational expectation equilibrium model.

[5]Andrei and Cujean (2017) argue that word-of-mouth communication is a plausible mechanism, as it produces several predictions that are supported by empirical evidence (Moskowitz et al., 2012; Hong et al., 2000).

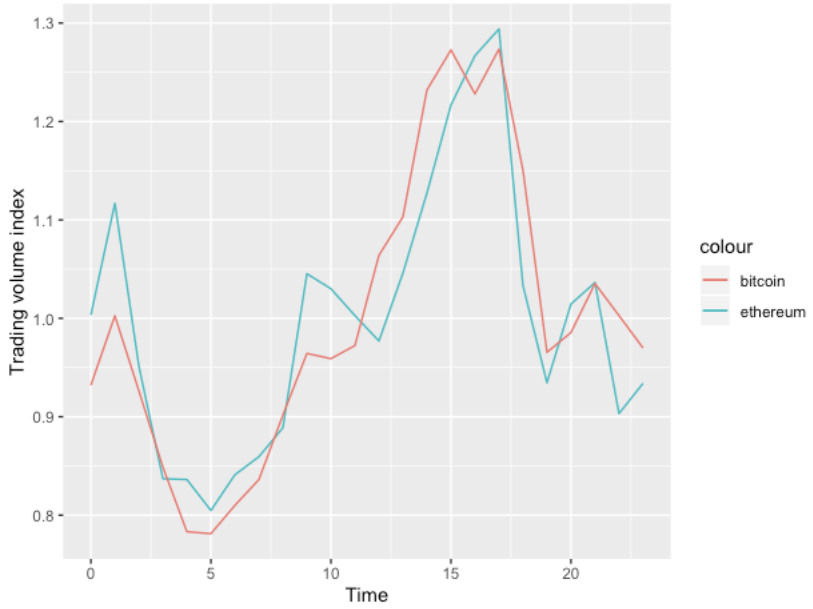
4.2 Two-way cuts on meeting intensity and trading hours

Next, I extend the analysis on Google trend index in Table 4.1 to double cuts. The method is very similar to the previous section except now we look at two additional subsamples cutting on working hours. It is natural to think that most of the investors will spend a specific time of the day regularly to make investment decisions. In addition, the feature of 24/7 continuous trading time for cryptocurrency market makes it possible for researchers to look more deeply into the variation of the momentum effect during a single day (24 hours). Taken together, it seems plausible to assume that meeting intensity will be larger when most of the investors are awake and active so they can meet often and diffuse information in a much faster way during the day than during the night. I further confirm this conjecture by plotting the total trading volumes during a single. As can be seen from Figure 4.2: it is true that most trades for both of the Bitcoin and Ethereum are executed between 9 a.m. and 5 p.m. (8 hours in total) rather than during the nighttime[6]. For naming convenience, I call the observations from the time interval 9 a.m. to 5 p.m. (8 hours in total) the daytime subsamples and the rest of the time interval (5 p.m. to 9 a.m.) observations the nighttime subsample. In the rest of the section I will show how the momentum effect varies in daytime subsample and nighttime subsample.

Table 4.2 presents the results of this approach. The short-term and long-term momentum effect are defined the same as in Table 4.1. However, the sorting becomes finer as I generate the subsamples by both of the AGT and the time interval for daytime/nighttime. The low meeting intensity and nighttime subsample (Low MI

[6]The time zone used for the trading data is the standard Greenwich Mean Time Zone (GMT)

Figure 4.2: Total trading volumes for Bitcoin and Ethereum in every hour from 0 to 23 (sample from the 31st of January 2017 to the 27th of June 2019)



Note: The trading volume index is calculated by first aggregating the every minute trading volume into hourly trading volume in the sample, then dividing the hourly trading volume by the 24 hour trading volume sample mean. The time zone for the trading data is the same as the Google trend data – Greenwich Mean Time Zone (GMT).

and Nighttime) includes the observations that meet two criteria simultaneously: from the time interval 5 p.m. to 9 a.m., and below the first quartile of the meeting intensity distribution. The low meeting intensity (Low MI) includes the observations from 0 to the first quartile of the meeting intensity distribution. The high meeting intensity (High MI) includes the observations above the third quartile of the meeting intensity distribution. The high meeting intensity and daytime subsample (Low MI and Daytime) includes the observations from the time interval 9 a.m. to 5 p.m., and above the third quartile of the meeting intensity distribution. I thus conjecture that the frequency at which agents exchange information each other will have such a rank: High AGT with Daytime > High AGT in general > Low AGT in general > Low AGT with Nighttime. Only the two extremes of the AGT spectrum is used because it is unclear how to compare the middle level combinations of the Google trend index and time interval (e.g., it is difficult to determine whether the meeting frequency is higher for Mid-low AGT with Nighttime or Mid-high AGT with Daytime). The second and third rows are listed here in order to easily compare the difference between Table 4.2 and Table 4.1.

As can be seen from Table 4.2, the basic results from Table 4.1 carry over. From the first column we can see that the short-term momentum effect increases as the meeting intensity increases. More importantly, we observe a significant reversal at the lowest meeting intensity subsamples (Low MI and Nighttime). This result fits the Andrei and Cujean (2017)'s prediction, which states that, even with a short lag, when the magnitude of the meeting intensity is very small (below the momentum threshold), the return will exhibit reversal[7].

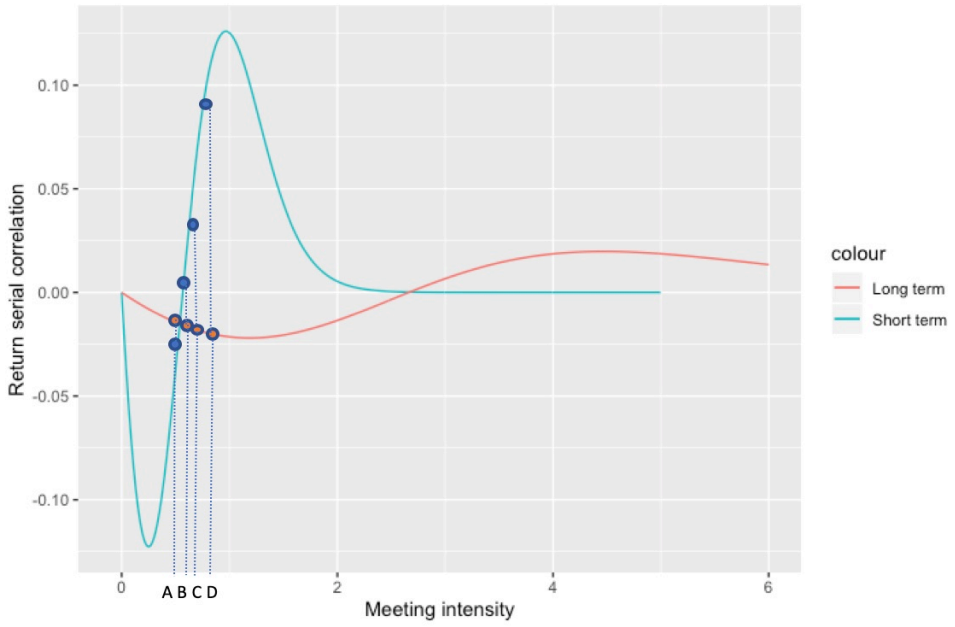
[7]When the meeting intensity is particularly low, the increase in the precision of private signals dominates, precision elasticity is less than one so that the percentage increase in average market precision exceeds the percentage increase in the precision of private information. Therefore the serial correlation of returns is negative. The detail derivations can be found in Theorem 3.1, Andrei and Cujean (2017).

Table 4.2: Momentum effect using Bitcoin returns and sorting by meeting intensity and time interval

	short-term	long-term	short - long	No. of Obs
Low MI and Nighttime	-0.005***	-0.002	-0.003***	191,000
	(0.002)	(0.002)	(0.000)	
Low MI	0.009**	-0.002	0.011***	288,266
	(0.002)	(0.002)	(0.000)	
High MI	0.066***	-0.011***	0.077***	288,266
	(0.002)	(0.002)	(0.000)	
High MI and Daytime	0.083***	-0.015***	0.098***	95,000
	(0.003)	(0.003)	(0.000)	
Highest - Lowest	0.088***	-0.013***	0.101***	
	(0.000)	(0.000)	(0.000)	

Note: This table lists the results of the estimates of $\beta_{s,h}$ in model 4.1 for different return lags. The first column lists the estimates of $\beta_{s,h}$ using lags of $h = 1$ minute for the short-term momentum effect. The second column lists the estimates of $\beta_{s,h}$ using lags of $h = 1$ hour for the long-term momentum effect. The last column shows the number of observations in respective subsamples. The rows report the estimates using an independent sort on the meeting intensity and time interval. The low meeting intensity and nighttime subsample (Low MI and Nighttime) includes the observations below the first quartile of the meeting intensity distribution, and from the time interval 5 p.m. to 9 a.m. including 16 hours in total. The low meeting intensity (Low MI) includes the observations from 0 to the first quartile of the meeting intensity distribution. The high meeting intensity (High MI) includes the observations above the third quartile of the meeting intensity distribution. The high meeting intensity and daytime subsample (Low MI and Daytime) includes the observations above the third quartile of the meeting intensity distribution, and from the time interval 9 a.m. to 5 p.m. including 8 hours in total. The third column and the last row are calculated using the two-sample t-test from the mean, standard deviation and sample size of the individual estimates. Standard errors are in parentheses. *** p < 0.001; ** p < 0.01; * p < 0.1

Figure 4.3: Simulated relationship between meeting intensity and serial correlation in returns with point estimates from Table 4.2



Note: The blue curve is drawn by the simulated points from the AC's model using lag $h = 1$ (short-term). The red curve is drawn by the simulated points using lag $h = 60$ (long-term). The calibration used is the same as in Andrei and Cujean (2017): $H = S = \Phi = 1$ and $\tau = 1/3$ where H is the precision of the security payoff, S is the precision of the signal distribution, Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock), τ is the investor risk tolerance. The four levels of meeting intensity A, B, C and D stand for the respective rows in Table 4.2: Low AGT + Nighttime, Low AGT in general , High AGT in general and High AGT + Daytime.

The second column again shows a negative relationship between the long-term momentum effect and the meeting intensity. It implies that as we move to progressively higher frequency meetings, the magnitude of the long-term reversal effect will rise, just like what we conclude from the second column in Table 4.1.

The third column shows that except for the ultra low meeting intensity subsample, the short-term momentum is statistically higher than the long-term momentum effect at 0.1% level of significance. In addition, the difference between the short-term and long-term momentum effect is increasing as the meeting intensity grows.

Similar to Figure 4.1, I also plot the serial correlation of returns as a function of the meeting intensity for different lags in Figure 4.3. The four levels of meeting intensities A, B, C and D here stand for the respective rows in Table 4.2: Low AGT with Nighttime, Low AGT in general, High AGT in general and High AGT with Daytime. We can again see a close match between the empirical estimates (dots) and the theoretical prediction (simulated curves).

As it can be seen from Figure 4.3, all the long-term (red points) return serial correlation estimates are below zero, which exhibit reversal. While all the short-term (blue points) return serial correlation estimates except the one with the lowest meeting intensity are above zero, therefore exhibit momentum. Furthermore, the empirical finding also matches what the model predicts about the difference between the short-term and long-term effect - as the meeting intensity rises, the difference will enlarge, which can also be seen from the third column of Table 4.2. Last but not least, the fifth row shows a significant positive effect of meeting intensity proxy on an asset's serial correlation at a short-term horizon and a negative effect at a long-term horizon, further prove the correctnesses of Hypothesis 1 and 2.

5 | Sensitivities

5.1 ETH results

This section replicates the exact same analysis as in Section 4.1, but applies to the second largest cryptocurrency - Ethereum. In other words, the s (currency pair parameter) in Eq. (5) is now changed from Bitcoin to Ethereum in order to check the robustness of the previous results. Table 5.1 reports the results of the estimates of $\beta_{s,h}$ using lags of $h = 1$ minute for the short-term effect and $h = 1$ hour for the long-term effect. Again, I use an independent sort to generate 4 subsamples by the meeting intensity proxy (AGT) - the low subsample which includes observations below the first quartile of the meeting intensity distribution, the mid-low subsample which includes the first quartile to the second quartile, the mid-high subsample which includes the second to third quartile and the high subsample which includes observations above the third quartile.

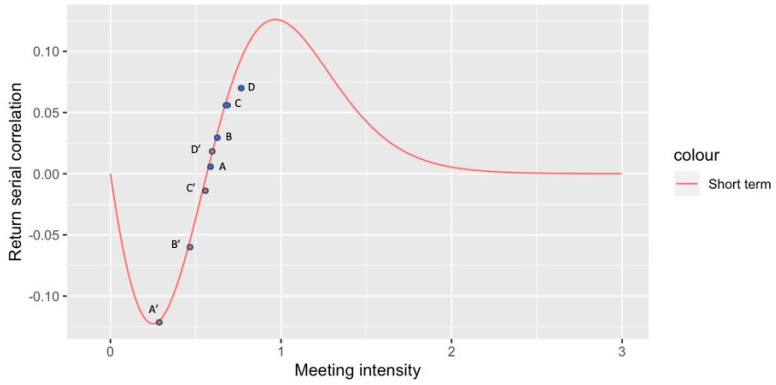
As it turns out, the use of second largest cryptocurrency does not significantly alter the previous key results. Both Hypothesis 1 and Hypothesis 2 stated in Chapter 4 still hold true. For Hypothesis 1, the last row of the first column shows the of the momentum effect with the large meeting intensity is significantly higher than the magnitude with the low meeting intensity at a short-term horizon. The difference of 0.178 between the two subsamples is highly statistically significant, with a p-value smaller than 0.1%. This result is even higher than the previous 0.057 difference in Table 4.1. For Hypothesis 2, the last row of the second column indicates a negative relationship between the autocorrelated returns and the meeting intensity at a long-term horizon. The difference in High - Low momentum is again highly statistically significant at -0.006, with a p-value smaller than 0.1%.

Table 5.1: Momentum effect using Ethereum returns and sorting by meeting intensity

	short-term	long-term	short - long	quartile
All	-0.06***	-0.002*	-0.058***	
	(0.001)	(0.001)	(0.000)	
Low MI	-0.158**	-0.004*	-0.154***	MI < 4.94
	(0.002)	(0.002)	(0.000)	
Mid-low MI	-0.062***	-0.006**	-0.056***	4.94 < MI < 7.54
	(0.002)	(0.002)	(0.000)	
Mid-high MI	-0.023***	0.001	-0.024***	7.54 < MI < 10.66
	(0.002)	(0.002)	(0.000)	
High MI	0.02***	-0.01***	0.03***	10.66 < MI
	(0.002)	(0.002)	(0.000)	
High - Low	0.178***	-0.006***	0.184***	
	(0.000)	(0.000)	(0.000)	

Note: This table lists the results of the estimates of $\beta_{s,h}$ in model 4.1 for different return lags. The first column lists the estimates of $\beta_{s,h}$ using lags of $h = 1$ minute for the short-term momentum effect. The second column lists the estimates of $\beta_{s,h}$ using lags of $h = 1$ hour for the long-term momentum effect. The last column shows the quartile of the meeting intensity (MI) distribution. The rows report the estimates using an independent sort on the meeting intensity (proxied by AGT). The low subsample (Low MI) includes observations from 0 to the first quartile of the meeting intensity distribution, the mid-low subsample includes the first quartile to the second quartile, the mid-high subsample includes the second to third quartile and the high subsample includes observations above the third quartile. The third column and the last row are calculated using the two-sample t-test from the mean, standard deviation and sample size of the individual estimates. The number of observations for each subsample is 288,266. Standard errors are in parentheses. *** p < 0.001; ** p < 0.01; * p < 0.1

Figure 5.1: Simulated relationship between meeting intensity and serial correlation in returns with point estimates from Table 4.1 and Table 5.1



Note: The curve is drawn by the simulated points using lag $h = 1$ (short-term). The calibration used is the same as in Andrei and Cujean (2017): $H = S = \Phi = 1$ and $\tau = 1/3$ where H is the precision of the security payoff, S is the precision of the signal distribution, Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock), τ is the investor risk tolerance. The four levels of meeting intensity A(A'), B(B'), C(C') and D(D') stand for the respective rows in Table 4.1 (Table 5.1).

Perhaps more interestingly, Table 5.1 shows a non-typical both short-term and long-term reversals in the full sample, as can be seen from the first row. In addition, the first three subsamples (Low MI, Mid-low MI and Mid-high MI) in the short-term column all report highly significant negative autocorrelations with p-values smaller than 0.1%. These results are quite different from the Bitcoin case in Table 4.1 where all the estimates in the first column are significant positive, implying a strong short-term momentum phenomenon. However, this seeming contradiction can be easily justified by using the Andrei and Cujean (2017)'s theory: Because Bitcoin is the most popular cryptocurrency, the magnitude of the overall meeting intensity for Bitcoin should be much larger than that for Ethereum. In fact, the ratio of Google trend index for keyword "bitcoin" and for "ethereum" is about 2:1 (information obtained on 1 Oct, 2019 from trends.google.com). Therefore the meeting intensity for Ethereum during the sample period may not exceed the momentum threshold. To better illustrate the comparison between different meeting intensity scales for "bitcoin" and "ethereum", I plot the Figure 5.1, where the meeting intensity scale for "ethereum" is reduced by approximately half from the original scale. The curve is drawn by the simulated points using lag $h = 1$ (short-term). The calibration used is again the same as in Andrei and Cujean (2017): $H = S = \Phi = 1$ and $\tau = 1/3$ where H is the precision of the security payoff, S is the precision of the signal distribution, Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock), τ is the investor risk tolerance. The four levels of meeting intensity A(A'), B(B'), C(C') and D(D') stand for the respective rows in Table 4.1 (Table 5.1). Again it is important to bear in mind that when comparing the cardinal value of the Poisson arrival rate (the actual meeting strength parameter in the model) and the ordinal rank from the Google index, care should be taken. This being stated, we can see a near fit between the empirical predictions (dots) and the theoretical projection (simulated

curves) after taking the relative meeting intensity scales of Bitcoin and Ethereum into account.

5.2 Different specifications of sampling frequency

This section applies the same analysis from Section 4.1 to different short-term and long-term sampling frequencies. More specifically, I examine the robustness of the results for the Hypothesis 1 in Section 4.1 using different specifications of short-term by using the 5, 10, and 30 minutes sampling frequencies. For the Hypothesis 2, I apply 3, 6 and 9 hours for the long-term specifications. Table 5.2 reports the results of the estimates of $\beta_{s,h}$ in model 4.1 using these lags for the short-term and long-term effects. The low subsample (Low MI) includes observations from 0 to the first quartile of the meeting intensity distribution, and the high subsample includes observations above the third quartile. The High-Low row is calculated using the two-sample t-test from the mean, standard deviation and sample size of the individual estimates.

As can be seen from the last row of Table 5.2, the previous results of testing Hypothesis 1 and 2 are both robust across different sampling frequencies. The first half of Table 5.2 (short-term) shows all highly positive High - Low momentum and the second half of Table 5.2 (short-term) shows all highly negative High - Low momentum, with p-values all smaller than 0.1%. The largest magnitudes in 1 minute (0.057) for short-term and 1 hour (-0.009) for long-term are consistent with the information inefficiency frequency findings documented in the recent literature (Aslan and Sensoy, 2019).

Table 5.2: Sampling frequency robustness check: momentum effect using Bitcoin returns and sorting by meeting intensity

	short-term			
	1 min	5 mins	10 mins	30 mins
Low MI	0.009***	-0.013***	0.001	0.003*
	(0.002)	(0.002)	(0.002)	(0.002)
High MI	0.066***	-0.011***	0.003*	0.006**
	(0.002)	(0.002)	(0.002)	(0.002)
High-Low	0.057***	0.002***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
	long-term			
	1 hr	3 hrs	6 hrs	9 hrs
Low MI	-0.002	0.000	0.002	-0.002*
	(0.002)	(0.002)	(0.002)	(0.002)
High MI	-0.011***	-0.004*	-0.001	-0.003*
	(0.002)	(0.002)	(0.002)	(0.002)
High-Low	-0.009***	-0.003***	-0.003***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)

Note: This table lists the results of the estimates of $\beta_{s,h}$ in model 4.1 for different Bitcoin return lags. The first and second rows (Low MI and High MI) report the estimates using an independent sort on the meeting intensity (proxied by AGT). The low subsample (Low MI) includes observations from 0 to the first quartile of the meeting intensity distribution, and the high subsample includes observations above the third quartile. The High-Low row is calculated using the two-sample t-test from the mean, standard deviation and sample size of the individual estimates. Standard errors are in parentheses. *** p < 0.001; ** p < 0.01; * p < 0.1

5.3 Regression approach

In this section, a different approach is taken to conduct the hypothesis testings. Within the most common terms, Hypothesis 1 (Hypothesis 2) is saying that the momentum effect is increasing (decreasing) with the meeting intensity for a short-term (long-term) lag. A simple way to test this would be to estimate a serial correlation coefficient for Bitcoin returns for the short-term and long-term lag respectively, and then regress this serial correlation coefficient on measures of the Google trend index.

More precisely, I first divide and regroup the full dataset from the 31st of January 2017 to the 27th of June 2019 into 819 days datasets (therefore 819 subgroups), where each subgroup contains 1,440 observations (60 minutes multiply 24 hours). Then for each day d, I regress the minute return $r^s_{t,d}$ for cryptocurrency pair s in time t on its return lagged h periods:

$$r^s_{t,d} = \alpha + \beta_{s,h,d} r^s_{t-h,d} + \epsilon^s_{t,d} \tag{5.1}$$

Where $\beta_{s,h,d}$ is the return serial correlation for pair s and lag h in day d, and $\epsilon^s_{t,d}$ is the statistical noise (e.g, measurement error). Just as before, the regressions are run using lags of $h = 1$ minute for the short-term effect and $h = 1$ hour (60 minutes) for the long-term effect.

After obtaining the estimates of $\beta_{s,h,d}$ for both cryptocurrency type s, lag h and day d, I perform the second regression , running $\beta_{s,h,d}$ on meeting intensity $MI_{s,d}$ (proxied by the daily adjusted Google Trends), which I call it **Model 1** (baseline

model):

$$\beta_{s,h,d} = \gamma_{s,h} MI_{s,d} + e_d^s \tag{5.2}$$

Where $\gamma_{s,h}$ is the impact of meeting intensity on momentum for pair s and lag h, and e_d^s is the statistical noise.

Similar to the two-way cuts in Section 4.2, in order to measure the impact of daytime/nighttime on momentum effect using the regression method, I construct a dummy variable: trading interval TI into Model 1, where $TI = 1$ if time is between 9 a.m. to 5 p.m. (8 hours in total) , otherwise $TI = 0$. Then I regress $\beta_{s,h,d}$ on meeting intensity $MI_{s,d}$ and $TI_{s,d}$, which I call it **Model 2**:

$$\beta_{s,h,d} = \gamma_{s,h} MI_{s,d} + \iota_{s,h} TI + e_d^s \tag{5.3}$$

Where $\gamma_{s,h}$ is the impact of meeting intensity on momentum for pair s and lag h, $\iota_{s,h}$ is the impact of trading interval on momentum for pair s and lag h, and e_d^s is the statistical noise. Notice here the full dataset is divided into 1,638 half days (819 * 2 subgroups) in order to construct the trading interval dummy variable.

Before turning to the regression results, It is useful to discuss how this general approach compares to what is done in Chapter 4. The main difference is that it imposes a more parametric structure, some of which may be unwarranted. For example, if there is a significant high-frequency (seconds) surge or plunge in meet-

Table 5.3: Momentum regression for Bitcoin and Ethereum

		Model 1		Model 2	
asset	var	short-term	long-term	short-term	long-term
BTC	MI	0.016***	-0.001**	0.016***	-0.001**
		(0.003)	(0.000)	(0.003)	(0.000)
	TI			0.019***	-0.003*
				(0.006)	(0.002)
ETH	MI	0.03***	-0.001*	0.03***	-0.001*
		(0.003)	(0.001)	(0.003)	(0.001)
	TI			0.009*	-0.003*
				(0.007)	(0.002)

Note: This table lists the results of the estimates in Model 1 and Model 2 for different return lags and different cryptocurrency pairs. Standard errors are in parentheses. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.1$

ing intensity, the less-structured method of the Chapter 4 will be more appropriate to handle. The offsetting benefit is that if the parametric form that I enforce with the regression is not too appropriate, our statistical capacity will be improved in those dimensions.

Table 5.3 summarizes the results. All the evidence points to a consistent positive effect of meeting intensity on an asset's serial correlation in the short-term and negative effect in the long-term. It is worth noting that even though the short-term positive effect of meeting intensity is strongly statistically positive with p-values smaller than 0.1% for both Bitcoin and Ethereum, the long-term negative effect of meeting intensity is not that strong. This is likely due to the fact that fewer observations are available to construct the long-term return serial correlation (since more observations are omitted in the long-term lag equation). For the trading interval effect, we can see from the Model 2 that there exists a statistically strong positive impact of trading interval (working time from 9 a.m. to 5 p.m.) to momentum at a short-term horizon and a negative impact at a long-term horizon. Again, the trading interval can be considered a good way to proxy meeting in-

tensity in addition to Google trend index because when most of the investors are awake and active from 9 a.m. to 5 p.m., they can meet more often and diffuse information in a more frequent way than during the night (5 p.m. to 9 a.m.). Therefore both Google Trends and trading interval regression results reassure the correctness of Hypothesis 1 and 2.

To sum up, Table 5.3 provides extra assurance around the robustness of our central results. Even with a very altered measurement approach, I obtain the same outcome for Hypothesis 1 and 2.

6 | Conclusion

As Fama (1998) states, one should not be too amazed if a newly generated model merely rationalize those existing phenomenons that it is deliberately designed to catch. Rather, the more useful test should be the "out-of-sample" one: The capacity to generate new hypotheses that are eventually proved true or false in future empirical research: "The over-riding question should always be: Does the new model produce coherent rejectable predictions . . .". Guided by this sentiment, this thesis tries to take one step further in the "out-of-sample" hypothesis direction.

Since the information percolation model of Andrei and Cujean (2017) was built for explaining the driving forces of both short-term momentum and long-term reversal, in accordance with Fama, it should be tested more on the basis of other predictions which have not been done in the past. Therefore, I choose two relatively simple and clear-cut such hypotheses: If momentum comes from information percolation and different investors own heterogeneous information, therefore perform divergent trading strategies, asset returns should display more positively autocorrelated returns when meeting intensity is high at a short-term horizon and display more negatively autocorrelated returns when meeting intensity is high at a long-term horizon.

Instead of repeating the aforementioned findings, at this time I think it should be sufficient to say that all of the evidence suggests the above hypotheses are true in the scope of the cryptocurrency market. However, it is not to declare that alternate interpretations of part or all of the findings can not be brought forward. If specific new interpretations are indeed offered, further rigorous work would have to be undertaken to verify the theory.

Bibliography

Abraham, J., Higdon, D., Nelson, J., and Ibarra, J. (2018). Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis. 1(3):22.

Andrei, D. and Cujean, J. (2017). Information percolation, momentum and reversal. *Journal of Financial Economics*, 123(3):617–645.

Aslan, A. and Sensoy, A. (2019). Intraday efficiency-frequency nexus in the cryptocurrency markets. *Finance Research Letters*.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3):307–343.

Cheah, E.-T. and Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130:32–36.

Conrad, J. and Kaul, G. (1998). An Anatomy of Trading Strategies. *The Review of Financial Studies*, 11(3):489–519.

Daian, P., Goldfeder, S., Kell, T., Li, Y., Zhao, X., Bentov, I., Breidenbach, L., and Juels, A. (2019). Flash Boys 2.0: Frontrunning, Transaction Reordering, and Consensus Instability in Decentralized Exchanges. *arXiv:1904.05234 [cs]*.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, 53(6):1839–1885.

Duffie, D. and Manso, G. (2007). Information Percolation in Large Markets. *American Economic Review*, 97(2):203–209.

Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, 17:81–91.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral. *Journal of Financial Economics*, 49(3):283–306.

Fama, E. F. and French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance*, 51(1):55–84.

Fang, F., Ventre, C., Basios, M., Kong, H., Kanthan, L., Li, L., Martinez-Regoband, D., and Wu, F. (2020). Cryptocurrency Trading: A Comprehensive Survey. *arXiv:2003.11352 [q-fin]*.

Fry, J. and Cheah, E.-T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47:343–352.

Garcia, D., Tessone, C. J., Mavrodiev, P., and Perony, N. (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of The Royal Society Interface*, 11(99):20140623.

Gkillas, K. and Katsiampa, P. (2018). An application of extreme value theory to cryptocurrencies. *Economics Letters*, 164.

Grossman and Stiglitz (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*, page 17.

Hau, Y.-S. and Kim, Y.-G. (2011). Why would online gamers share their innovation-conducive knowledge in the online game user community? Integrating individual motivations and social capital perspectives. *Computers in Human Behavior*, 27:956–970.

Hong, H., Lim, T., and Stein, J. C. (2000). Bad News Travels Slowly: Size,

Analyst Coverage, and the Profitability of Momentum Strategies. *The Journal of Finance*, page 31.

Hong, H. and Stein, J. C. (1999). A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *The Journal of Finance*, page 42.

Jegadeesh, N. and Titman, S. (2001). Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *The Journal of Finance*, 56(2):699–720.

Katsiampa, P., Corbet, S., and Lucey, B. (2019). High frequency volatility co-movements in cryptocurrency markets. *Journal of International Financial Markets, Institutions and Money*, 62:35–52.

Koutmos, D. (2018). Return and volatility spillovers among cryptocurrencies. *Economics Letters*, 173:122–127.

Lewis, M. (2014). *Flash Boys: Revolte an der Wall Street*. Campus Verlag.

Makarov, I. and Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, 135(2):293–319.

Moskowitz, T. J., Ooi, Y. H., and Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, 104(2):228–250.

Panzarasa, P., Opsahl, T., and Carley, K. M. (2009). Patterns and dynamics of users' behavior and interaction: Network analysis of an online community. *Journal of the American Society for Information Science and Technology*, 60(5):911–932.

Philippas, D. (2019). Media Attention and Bitcoin Prices. *SSRN Electronic Journal*.

Phillip, A., Chan, J. S. K., and Peiris, S. (2018). A new look at Cryptocurrencies. *Economics Letters*, 163(C):6–9.

Weber, B. (2016). Bitcoin and the legitimacy crisis of money. *Cambridge Journal of Economics*, 40(1):17–41.

Yermack, D. (2013). Is Bitcoin a Real Currency? An economic appraisal. Working Paper 19747, National Bureau of Economic Research.

Appendix

Appendix A

Cryptocurrency Price

The historical prices of Bitcoin and Ethereum are collected every minute via the Bitfinex API and placed into text files to create a price history. The Python code can be found at https://github.com/jianwang0212.

Google Trends Adjustment

Google trends data is an unbiased sample of search data starting from 2004. Instead of providing search volumes, Google is using a search volume index (SVI). The search volume index is calculated by dividing each data point by the total searches within a geographic region and time range. The numbers are then scaled between 0 and 100 on a search term's proportion to all searches on all topics. When trends data is queried for a period of longer than 90 days the SVI returned are aggregated at a weekly level. In order to compare these SVIs across periods

and adjustment has to be made. For our research we used the method detailed by Erik Johansson[1].

The method has four primary steps. First, collect all of the daily SVI data you need in 90 day increments and combine them into a single increment covering the entire time period of interest. Second, line up the data for the same entire time period, but aggregated at a weekly level to get the weekly SVI. Determine an adjustment factor which is done by dividing the weekly SVI with the daily SVI value where the dates overlap. Finally, multiple the daily SVI values by the adjustment factor. In cases when the SVI was less than 1, the value was returned by the Google Trends query as < 1. To allow for an adjustment calculation we changed that value to 0.5. Google does not provide any more information on what the specific value was, so the halfway value of 0.5 was used as a substitute.

Appendix B

Momentum condition equivalence

To achieve a nice intuition of the momentum condition ($m_{t-l} > 0$), I will redefine this condition in terms of the precision elasticity (ϵ_t) and show an equivalence

[1] http://erikjohansson.blogspot.com/2014/12/creating-daily-search-volume-data-from_8.html

between $m_t > 0$ and $\epsilon_t > 1$ where $\epsilon_t \equiv \frac{(K_{t+1}-K_t)/K_t}{(\sum_{k=0}^{t+1} S\Omega_t - \sum_{k=0}^{t} S\Omega_t)/\sum_{k=0}^{t} S\Omega_t}$:

$$\begin{aligned} m_t &> 0 \\ \sum_{k=0}^{t} S\Omega_k - \frac{S\Omega_{t+1}}{(K_{t+1}-K_t)/K_t} &> 0 \end{aligned} \tag{6.1}$$

$$\begin{aligned} \frac{(K_{t+1}-K_t)/K_t}{S\Omega_{t+1}/\sum_{k=0}^{t} S\Omega_t} &> 1 \\ \frac{(K_{t+1}-K_t)/K_t}{(\sum_{k=0}^{t+1} S\Omega_t - \sum_{k=0}^{t} S\Omega_t)/\sum_{k=0}^{t} S\Omega_t} &> 1 \end{aligned} \tag{6.2}$$

Relationship between return serial correlation and the meeting intensity

For simplicity, I only consider an economy with T= 3 trading dates, indexed by $t = 0, 1, 2$ and a final liquidation date, $T = 3$. The risky security with payoff $\tilde{U}$ realized at the liquidation date. Denote returns between t and $t-1$ as $\tilde{R}_t = \tilde{P}_t - \tilde{P}_{t-1}$. Fix the lag $l = 1$.

When $t = 1$

$$\begin{aligned} \mathbb{E}[\tilde{P}_2 - \tilde{P}_1|\mathcal{F}_1^r] &= \frac{K_2 - K_1}{K_2 K_1^c}[m_0(\tilde{P}_1 - \tilde{P}_0) + m_{-1}(\tilde{P}_0 - \tilde{P}_{-1})] \\ \mathbb{E}[\tilde{R}_2|\mathcal{F}_1^r] &= \frac{K_2 - K_1}{K_2 K_1^c}[m_0(\tilde{R}_1) + m_{-1}(\tilde{R}_0)] \\ &= \frac{K_2 - K_1}{K_2 K_1^c} m_0 * \tilde{R}_1 \\ &= \rho * \tilde{R}_1 \end{aligned}$$

where

$$\begin{aligned}\rho &= \frac{K_2 - K_1}{K_2 K_1^c} m_0 \\ &= \frac{K_2 - K_1}{K_2 K_1^c} * (S\Omega_0 - \frac{S\Omega_1}{(K_1 - K_0)/K_0}) \\ &= \frac{S(K_1 - K_0 - \Omega_1 K_0)(K_2 - K_1)}{(K_1 - K_0) K_2 K_1^c} \\ &= f(\lambda, H, S, \Phi, \tau)\end{aligned}$$

is a function of parameters $\lambda, H, S, \Phi, \tau$.

where λ is the Poisson distribution parameter (meeting intensity),

H is the precision of the security payoff,

S is the precision of the signal distribution,

Φ is the precision of the aggregate per capita supply of the risky asset (net supply shock),

τ is the investor risk tolerance.